ARTICLE   OPEN

Optimal frequency measurements with quantum probes

Simon Schmitt[1], Tuvia Gefen[2], Daniel Louzon[1,2], Christian Osterkamp[1], Nicolas Staudenmaier[1], Johannes Lang[1], Matthew Markham[3], Alex Retzker [2], Liam P. McGuinness [1,4✉] and Fedor Jelezko[1,5]

Precise frequency measurements are important in applications ranging from navigation and imaging to computation and communication. Here we outline the optimal quantum strategies for frequency discrimination and estimation in the context of quantum spectroscopy, and we compare the effectiveness of different readout strategies. Using a single NV center in diamond, we implement the optimal frequency discrimination protocol to discriminate two frequencies separated by 2 kHz with a single 44 µs measurement, a factor of ten below the Fourier limit. For frequency estimation, we achieve a frequency sensitivity of 1.6 µHz/Hz² for a 1.7 µT amplitude signal, which is within a factor of 2 from the quantum limit. Our results are foundational for discrimination and estimation problems in nanoscale nuclear magnetic resonance spectroscopy.

INTRODUCTION

Quantum sensing uses platforms such as photons, ions, solid-state defects, and their quantum properties as resources to estimate physical quantities as precisely as possible[1,2]. The performance depends on the sensing and readout protocols, which should optimize the ratio of the sensor response for the parameter of interest[3,4] to readout noise[5]. Thus, finding optimal protocols is crucial to enabling efficient estimation. One of the major applications of quantum sensing is nanoscale nuclear magnetic resonance (NMR) spectroscopy in which a nanoscale quantum sensor replaces the macroscopic inductive coil and interacts with a sample of nuclear spins[6–9]. Pioneering work with the nitrogen-vacancy (NV) center in diamond[10,11], has demonstrated nanoscale spatial resolution[6,7,12–14] with single spin sensitivity[9]. Understanding and realizing the limits of quantum measurements is particularly important in spectroscopy wherein frequency encodes energy, spatial, and structural information.

Here we determine the quantum limit for discriminating known frequencies, develop a protocol that saturates it, and prove that this protocol achieves the minimal possible error probability as a function of time. After deriving the theoretical limits, we experimentally demonstrate this protocol by discriminating two known frequencies separated by 2 kHz, with a single 44 µs (Fourier limit $1/T \sim 23$ kHz) coherent measurement. We extend our studies by explicitly analyzing the influence of imperfect readout of our sensor qubit and perform a detailed comparison between two readout strategies. We show that these results can also be applied to find the optimal protocol for frequency estimation[15–19] and we use this protocol to experimentally estimate the value of a single unknown frequency. The relevance of the estimation protocol for realistic nano-NMR scenarios and the implications of the imperfect NV readout is further analyzed. Our results are important when one is faced with a decision on how to allocate finite resources to construct better sensors.

RESULTS

Optimal frequency discrimination using a quantum probe

As a diagnostic tool, NMR can be used to answer "yes–no" questions such as whether a certain toxin or metabolite is present in the sample. As sketched in Fig. 1a, the task is then to discriminate between two known spectra based on their frequency components. We define discrimination error as the error to decide on the wrong spectrum and our goal is to obtain a minimal discrimination error or equivalently a minimal discrimination time. A typical method compromises sampling the signal with consecutive, synchronized measurements and correlating the individual outcomes, e.g., by applying Fourier analysis, to obtain a spectrum. For sufficient recording time, the resolution can be high enough for an almost error-free discrimination. To illustrate, consider a simplified problem in which one wishes to discriminate between two spectra, each containing only a single frequency (ω_1 or ω_2) with the same amplitude B (in units of angular frequency). Naively, the method described above is Fourier limited, i.e., the time required for discrimination is $T = \frac{1}{|\omega_\Delta|}$, where $\omega_\Delta = \omega_2 - \omega_1$. Using a more sophisticated data analysis, such as Bayesian interference or machine learning, which can be applied for known B, the discrimination time lowers to $T \sim \frac{1}{B^{2/3}\omega_\Delta^{2/3}}$ (see Supplementary Note 1 and ref. [20]) However, we show in the following that, given a sufficient coherence time, the discrimination time can be further reduced to:

$$T = \frac{\pi}{2\sqrt{B}\sqrt{|\omega_\Delta|}} \qquad (1)$$

The key idea is to drive the sensor such that the angle between the states, $|\psi(\omega_1)\rangle$ and $|\psi(\omega_2)\rangle$, is maximal, to guarantee a minimal error probability. Once orthogonality is achieved, quantum projection noise can theoretically be eliminated by measuring in the appropriate basis and it is possible to determine, with a certainty limited by readout fidelity, which frequency is present. Assuming a perfect readout, we show that this method is in general optimal, even if orthogonality cannot be achieved.

[1]Institute of Quantum Optics, Ulm University, Ulm, Germany. [2]Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem, Israel. [3]Element Six, Didcot, UK. [4]Laser Physics Centre, Research School of Physics, Australian National University, Canberra, ACT, Australia. [5]Center for Integrated Quantum Science and Technology (IQST), Ulm University, Ulm, Germany. ✉email: liam.mcguinness@anu.edu.au
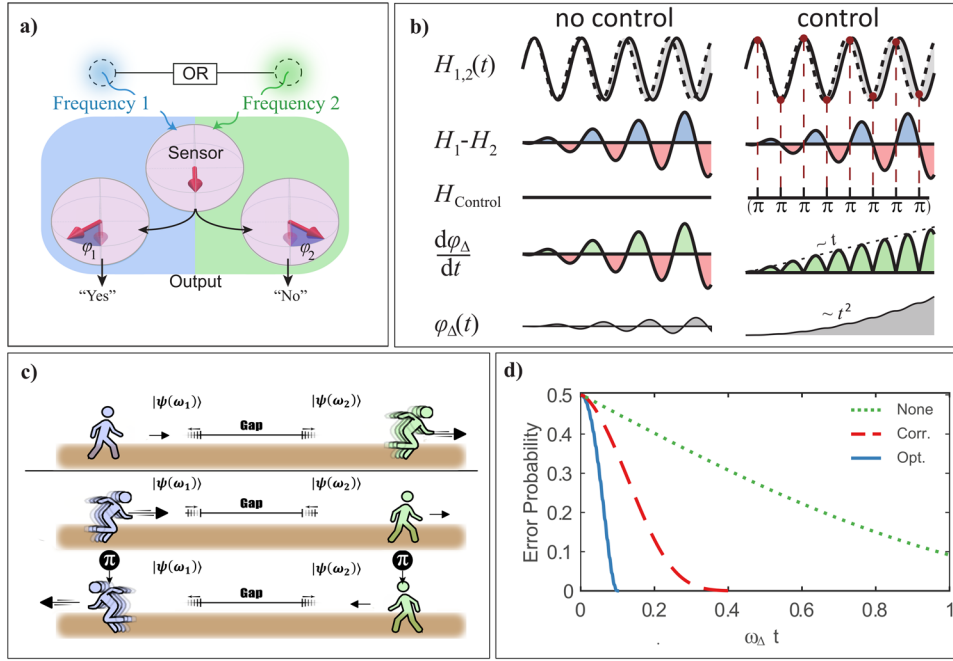
Fig. 1 Quantum-frequency discrimination. a The sensor state is tailored to depend on each frequency so that the sensor is driven in orthogonal directions. Readout of the state provides a "Yes/No" answer. The task is to correctly identify the frequency with highest fidelity in the shortest time. **b** Phase accumulation for frequency discrimination. For each Hamiltonian $\mathcal{H}_i(t) = B \sin(\omega_i t + \theta)\sigma_Z = H_i(t)\sigma_Z$, the sensor accumulates a different phase. Left: without control, the phase difference φ_Δ oscillates and increases only slowly. Right: optimal control in this scenario implies applying π-pulses whenever the sign of H_1-H_2 changes, leading to a monotone phase increase with t^2. **c** Geometrical picture of the optimal protocol. The states of the sensor can be thought of as two runners, where the aim is to maximize the gap between them, which is equivalent to the angle between the states. The speed of each runner is proportional to $H_i(t)$. As soon as the initially slower runner becomes faster, we change their direction of motion to ensure an increasing gap. Note that the sketch just demonstrates the idea and the runners actually move on circular orbits. **d** Error probability as a function of time for three different strategies: optimal control and a single measurement (solid blue), this strategy is optimal and sets the fundamental error limit (see Eqs. 3 and 4). Correlated measurements (dashed red), in this illustration measurement is applied every $\frac{2\pi}{\omega_1+\omega_2}$. No control (dotted green). More details on this comparison are found in Supplementary Note 1.

Given free evolution of the sensor, the discrimination time is $T \sim \frac{\omega}{B\omega_\Delta}$ (see Supplementary Note 1), which is basically the same scaling as the Fourier limit. As sketched in Fig. 1b, the phases $\varphi(\omega_1)$ and $\varphi(\omega_2)$, accumulated by the sensor under ω_1 and ω_2, move apart and get closer to each other such that their difference $\varphi_\Delta = \varphi(\omega_1) - \varphi(\omega_2)$ oscillates and only slowly increases. However, T can be significantly reduced by applying a suitable control, which is also shown in Fig. 1b: whenever the states start to get closer to each other (φ_Δ reduces), a control π-pulse can be used to change the direction of motion such that their distance increases instead. Formally, the distance between two states can be described by the angle $\varphi_\Delta/2 \underset{\text{def}}{=} a(\omega_1, \omega_2)$, which evolves as:

$$\frac{da(\omega_1, \omega_2)}{dt} = (H_1(t) - H_2(t)) \tag{2}$$

where $\mathcal{H}_i(t) = B \sin(\omega_i t + \theta)\sigma_Z = H_i(t)\sigma_Z$ are the corresponding signal Hamiltonians, θ is the initial signal phase, and σ_Z is the Pauli spin-z operator. As a consequence, maximizing $a(\omega_1, \omega_2)$ implies applying a π-pulse whenever $H_1(t) - H_2(t)$ changes sign and it follows (see Supplementary Note 1):

$$a(\omega_1, \omega_2) = \int_0^t \frac{\mu_{max} - \mu_{min}}{2} dt' \approx B\frac{2}{\pi}\omega_\Delta t^2 \tag{3}$$

where μ_{max} and μ_{min} are the maximal and minimal eigenvalues of $\mathcal{H}_1 - \mathcal{H}_2$. Application of a series of π-pulses with a spacing $\frac{2\pi}{\omega_1+\omega_2}$ constitutes optimal control in this scenario. As orthogonality requires $a(\omega_1, \omega_2) = \pi/2$, the minimal discrimination time is $T_{opt} = \frac{\pi}{2\sqrt{B}\sqrt{|\omega_\Delta|}}$.

For an illustrative explanation, one can think of $|\psi(\omega_1)\rangle$, $|\psi(\omega_2)\rangle$ as two runners, where the goal is to maximize the gap between them (see Fig. 1c). As long as the same runner (suppose $|\psi(\omega_2)\rangle$) is faster (speed proportional to the current amplitude), their separation gets larger as desired. However, once $|\psi(\omega_1)\rangle$ starts to be faster than $|\psi(\omega_2)\rangle$, we prevent a reduction of the gap by flipping the direction they run (equivalent to a π-pulse).

The phase acceleration in Eq. (4) coincides with the fundamental speed limit derived in ref. [21] and this implies a minimal discrimination error for every t. Hereafter, for all error analysis, we assume that the prior probability for both frequencies is 1/2, namely symmetric hypothesis testing. Given any two Hamiltonians $\mathcal{H}_1(t)$ and $\mathcal{H}_2(t)$, we prove in Supplementary Note 1 that the error probability when distinguishing between these Hamiltonians, optimized over all possible strategies, is lower bounded by:

$$\epsilon \geq \begin{cases} 0.5(1 - \sin(a_{max})) & a_{max} \leq \frac{\pi}{2} \\ 0 & a_{max} > \frac{\pi}{2} \end{cases} \tag{4}$$

where $a_{max} = \int_0^t \frac{\mu_{max} - \mu_{min}}{2} dt'$. This lower bound can always be saturated with a suitable control. Hence, even if $a_{max} \leq \frac{\pi}{2}$ at t, the strategy that minimizes the error probability is to apply the above discussed control and perform a measurement at t, which implies that even multiple correlated measurements at times shorter than t[19,22] cannot beat this fundamental limit, as shown in Fig. 1d (see Supplementary Note 1).

Experiments

The experiments are performed using a single NV center in ultrapure diamond (Fig. 2a). Here we use a 400 G magnetic field,
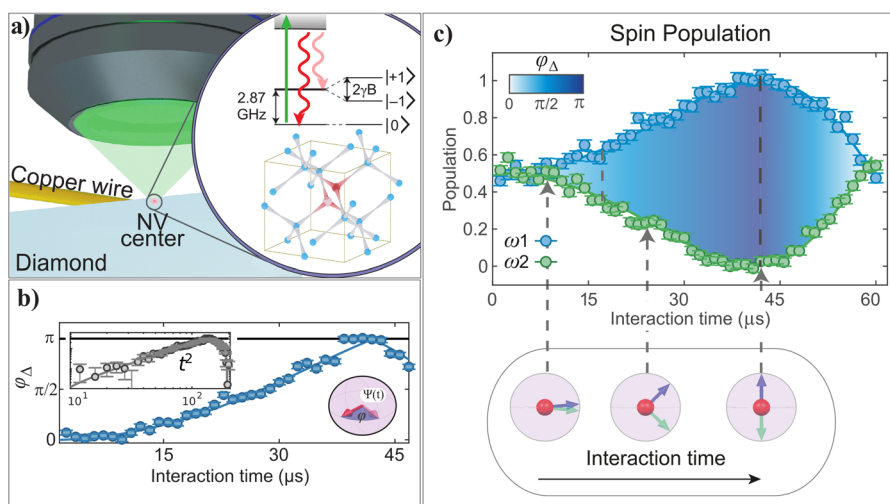
Fig. 2 **Quantum-frequency discrimination. a** Experimental setup: the spin state of a single nitrogen-vacancy (NV) center in diamond responds to magnetic fields by a Zeeman shift of the spin levels, which can be optically initialized and readout with a confocal microscope. **b** Experimentally measured quantum phase difference φ_Δ as a function of coherent interaction time and as log–log plot (inset). Solid lines are a fit to $\varphi_\Delta = \frac{2}{\pi} B \omega_\Delta t^2$, where $\omega_\Delta = |\omega_1 - \omega_2| = (2\pi) \cdot 2$ kHz. **c** Measured spin population as a function of coherent interaction time. The probability to be in state $|0\rangle$ is: $P(|0\rangle |\omega_1) = \frac{1}{2} - \frac{1}{2} \sin\left(\frac{B\omega_\Delta t^2}{\pi}\right)$, $P(|0\rangle |\omega_2) = \frac{1}{2} + \frac{1}{2} \sin\left(\frac{B\omega_\Delta t^2}{\pi}\right)$. For $\varphi_\Delta = \pi$, the sensor is driven to one of two orthogonal eigenstates. All error bars correspond to SD of several independent repetitions.

aligned along the NV symmetry axis, to lift the degeneracy of the three ground spin states and we use two of these states (here denoted as $|0\rangle$, $|1\rangle$) as a qubit. We initialize our qubit into a coherent superposition state and map the sensor phase $\varphi = 2\alpha$ (compare Eq. (3)), acquired during interaction with the control sequence and the signals, into a population difference (such that the probability for the sensor to be projected to $|0\rangle$ reads $P = 0.5 \cdot (1 - \sin(\varphi))$, which is subsequently read out optically.

In Fig. 2b, φ_Δ is plotted, when a magnetic field of frequency $\omega_1 = (2\pi) \cdot 0.999$ MHz or $\omega_2 = (2\pi) \cdot 1.001$ MHz, in the range of frequencies typical for nano-NMR experiments with NV centers and was measured with an XY8-N sequence with an interpulse spacing of 500 ns. The interaction time was extended by increasing the pulse number. A logarithmic plot of φ_Δ shows the expected t^2 increase until 44 μs when $\varphi_\Delta = \pi$, at which time the quantum sensor has been evolved into one of two orthogonal states, as $|P(\omega_1) - P(\omega_2)| = 1$. By using a 90° phase-shifted $\frac{\pi}{2} -$ pulse (compared to the initialization basis), the sensor phase can be mapped into a population difference and the resulting NV population as a function of interaction time is plotted (Fig. 2c). In Supplementary Note 4, we show that this balanced readout is also optimal. Experimental data for the dependency of the discrimination time on ω_Δ and B are provided in the Supplementary Note 5. If the coherence of the sensor is not sufficient to achieve orthogonality, the described protocol still remains optimal (see detailed discussion in Supplementary Note 1).

Complex frequency discrimination
This method can be extended for complex spectra consisting of more than one frequency or amplitude component. Consider a Hamiltonian that takes the general form $\mathcal{H}_{1,2} = \int f_{1,2}(\omega) \sin(\omega t) d\omega \, \sigma_Z = H_{1,2} \sigma_Z$ such that the two frequency distribution functions $f_{1,2}(\omega)$ are known, but we do not know which of the Hamiltonians, $\mathcal{H}_1$ or $\mathcal{H}_2$, is present. As an illustrative example, we consider the ^{13}C nuclear magnetic spectrum produced by either a sample of ethanol or propanol, both of which contain two chemically distinct carbon groups (Fig. 3a, b). A similar speed bound for discrimination applies to this case (see Supplementary Eq. 1) and is achievable with an analogous protocol. We can again define $H_\Delta = H_1 - H_2$, where optimal
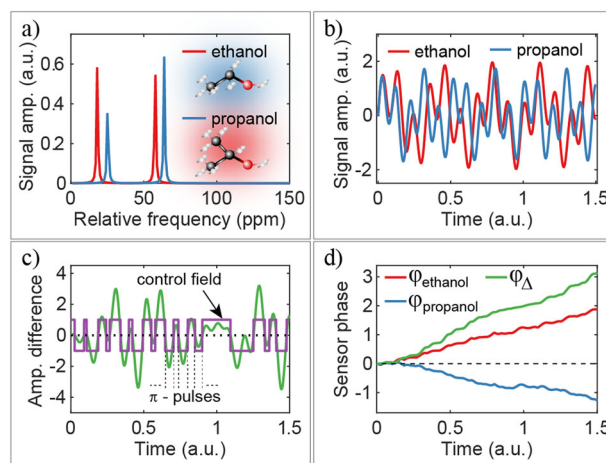

Fig. 3 **Optimal frequency discrimination of complex signals. a, b** Calculated magnetic field experienced by the NV center exposed to a solution of ethanol (red) or propanol (blue) in frequency (**a**) and time (**b**) domains, where a $\pi/2$ pulse is applied to the thermally polarized sample at zero time. **c** Calculated optimal control sequence of π-pulses applied to the NV sensor (purple) overlayed on the signal difference (green). **d** Sensor phase accumulation in presence of ethanol (red), propanol (blue), and the phase difference (green), showing quadratic scaling at short times.

control is achieved by applying π-pulses whenever H_Δ switches sign (Fig. 3c). Importantly, for NMR detection of a polarized sample, the initial signal phase is known, as it is defined by a $\pi/2$-pulse applied to the sample. As a result, a sensor phase difference of π can be tailored to perform optimal discrimination. Of note, the sensor phase difference also increases according to t^2, until the signals become completely out of phase with each other.

Implications of imperfect readout
Above, we showed that in principle, it is possible to discriminate two frequency distributions within a single measurement by eliminating projection noise. As perfect measurements are not possible in an experimental realization, additional readout noise
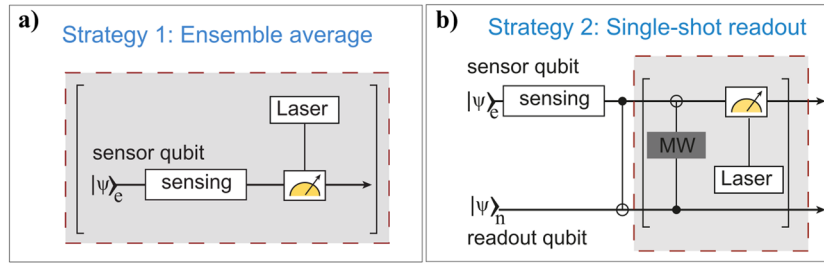
Fig. 4 Experimental realization of the two readout strategies—ensemble averaging and SSR. a Ensemble averaging. The complete sequence of sensor initialization, sensing period, and optical readout is performed N_{ens} times. **b** Single-shot readout (SSR). After one sensing duration, the sensor state is mapped onto the ancilla readout qubit (^{13}C nucleus). Optical readout of the ancilla spin is weakly perturbing, allowing its population to be probed N_{RR} times with a quantum non-demolition measurement.

has to be taken into account. For optical readout of NV centers, e.g., photon shot noise has to be considered. As a result, noise analysis is important to obtain a meaningful discrimination error. The readout of the NV center is performed by detecting the spin-dependent fluorescence emitted during a laser pulse (see "Methods"). The recorded photons are well described by a Poisson distribution with an average photon number of $\lambda_0 = 0.084$ ($\lambda_1 = 0.07$) for the $|0\rangle$ ($|1\rangle$) spin state (compare Supplementary Note 2). With this small contrast, the error of a single measurement, assuming again a symmetric hypothesis testing, is ~0.49, even for orthogonal states (see Supplementary Note 4), i.e., we have a nearly 50–50 chance to assign the frequency incorrectly.

The probability to make an incorrect decision can be reduced in two ways: first, by increasing the number of measurements, which we refer to as ensemble averaging. Then, the difference in the number of emitted photons increases, which leads to better discriminability. The second approach is based on improving the readout process itself by increasing the contrast between the states. This can be achieved by introducing an ancilla qubit, which acts as a quantum memory and stores the state of the NV center[23,24]. This method is usually referred to as single-shot readout (SSR). We experimentally implement these two methods and compare their performance using the error probability as figure of merit. We show that by benchmarking against the number of detected photons, ensemble averaging always performs at least as good as SSR, but when compared in terms of measurement time or number of coherent interaction periods, SSR has particular advantages.

Temporal ensemble averaging was performed using a 1.35 NA oil objective, to collect the NV fluorescence and repeating the sensing and readout N_{ens} times (compare Fig. 4a). Hence, given the orthogonal states and N_{ens} repetitions of the measurement, the discrimination between the states reduces to discrimination between two Poissonian distributions: $Poi(N_{ens}\lambda_0)$ and $Poi(N_{ens}\lambda_1)$. Assuming small contrast, namely $\frac{\lambda_0}{\lambda_1}$ is close to 1, we observe that the error probability scales as $\exp\left(-N_{ens}0.5\left(\sqrt{\lambda_1} - \sqrt{\lambda_0}\right)^2\right)$ (Supplementary Note 4). For non-orthogonal states, the error probability also decays exponentially with N_{ens} to 0 and the error exponent is $\sim \frac{(\lambda_0 - \lambda_1)^2 \sin(\alpha)^2}{4(\lambda_0 + \lambda_1)}$ (Supplementary Note 4), where α is the angle between the states. Hence, given that the duration of a single ensemble average is τ (including the overheads for readout and initialization), the discrimination time with this strategy goes as $\sim \tau \frac{4(\lambda_0 + \lambda_1)}{(\lambda_0 - \lambda_1)^2 \sin(\alpha)^2}$

SSR is implemented by using a weakly coupled ^{13}C spin that forms a memory qubit and allows repetitive readout of the NV state with a quantum non-demolition (QND) experiment[25] (compare Fig. 4b and Supplementary Note 2). The ancilla qubit allows more photons to be scattered before its state is destroyed and here we perform N_{RR} repetitive readouts after one sensing

time. The larger N_{RR}, the closer we get to a perfect quantum measurement, limited ultimately by the lifetime of the ancilla. For $N_{RR} = 10^4$, the photon statistics manifest digital step jumps, which allows high-fidelity readout of the nuclear spin state (see Supplementary Note 2). The discrimination between orthogonal states using this method is, in theory, a discrimination between $Poi(N_{RR}\lambda_0)$ and $Poi(N_{RR}\lambda_1)$, and thus given as $N_{RR} = N_{ens}$, precisely the same as for ensemble averaging. Then, the only difference is the time required to perform readout of the nuclear spin and that SSR requires only one sensing period, which is discussed in more detail below. For non-orthogonal states however, this strategy is inferior as the error probability does not reduce to 0 in the limit of large N_{RR} due to quantum projection noise. This additional uncertainty inherent to quantum measurements results in a scaling as $0.5(1 - |\sin(\alpha)|)$ (see Supplementary Note 1).

We experimentally investigated the performance of both strategies by performing ensemble averaging and SSR for discrimination of the same signals described earlier. The error as a function of different parameters is plotted in Fig. 5a–e. In Fig. 5a, it is shown that for any angle α between the states (or equivalently, for any coherent interaction period), ensemble averaging performs better than SSR. However, in contrast to the theoretical expectation, we see that ensemble averaging yields a lower error than SSR also for orthogonal states (44 μs, $\varphi_\Delta = \pi$). The reason for this is that, although the nuclear state can be readout with high fidelity in our experiments, imperfect initialization of the NV center into the correct charge state and the finite ancilla T_1 lifetime, limits the sensor readout fidelity to 0.8 (see Supplementary Note 2 and refs. [26,27]). In Fig. 5a (red line), we plot the expected error achievable with SSR under perfect sensor initialization and control, which demonstrates the equivalence between the two readout methods when orthogonality is achieved.

Figure 5b shows the discrimination error as a function of the number of photons detected with each readout technique, for a fixed interaction time of 44 μs ($\Delta\varphi = \pi$, dotted line in Fig. 5a). As the sensor is driven to an eigenstate, each photon should convey the same information about which frequency is present regardless of whether it was obtained from ensemble averaging or SSR. We observe that imperfect initialization reduces the information carried by each photon using SSR, resulting in the error converging to an offset of 0.2. By using charge-state detection to improve the sensor initialization, it is expected that a convergence in the two readout strategies would be observed.

However, SSR may still be useful if the total measurement time is taken into account. For ensemble averaging, every repetition cycle requires a duration of initialization (t_{init}), interaction (t), and readout (t_{read}); thus, the total measurement time is $T_{ens} = N_{ens}(t_{init} + t + t_{read})$. For our experiments, $t_{init} + t_{read} \approx 1.5$ μs. For SSR, only one initialization and interaction period are required, while readout is performed N_{RR} times. The initialization and readout time for SSR are different to ensemble averaging, however, as manipulation of the ancilla is required; thus, the total
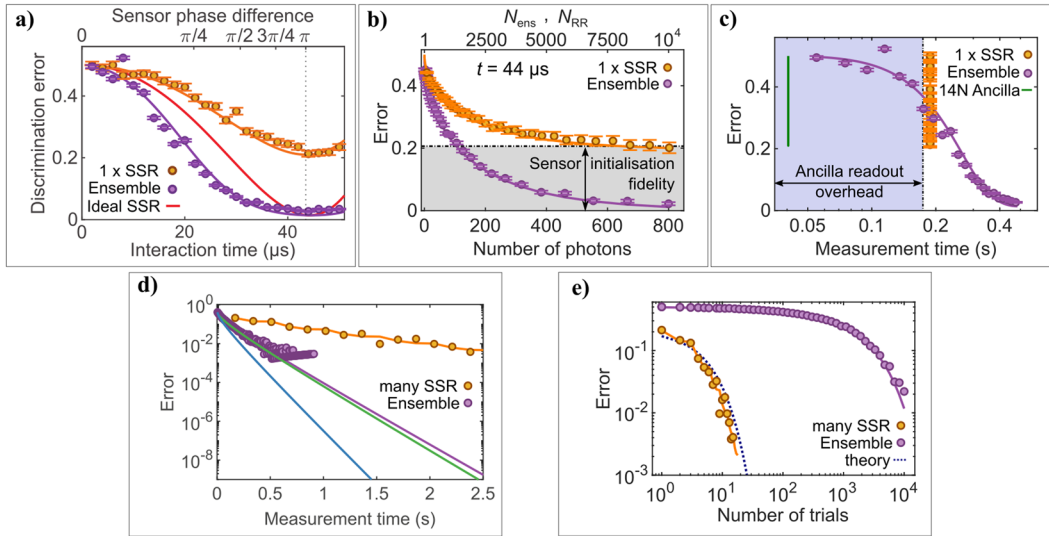
Fig. 5 Performance of quantum-frequency discrimination as a function of available resources for $\omega_1 = (2\pi) \cdot 0.999$ MHz, $\omega_2 = (2\pi) \cdot 1.001$ MHz. a Experimental frequency discrimination error as a function of coherent interaction time. Data points are for sensor readout using a single SSR measurement (orange, $N_{RR} = 10^4$) and ensemble averaging (lilac, $N_{ens} = 10^4$). Fits are from Supplementary Eq. 89 and 91. The ideal SSR curve corresponds to perfect sensor initialization and mapping to ancilla. **b** Experimental frequency discrimination error as a function of detected photons for a single SSR measurement (orange) and ensemble averaging (lilac), for an interaction time of 44 μs ($\varphi_\Delta = \pi$). Imperfect sensor initialization limits the error achievable with a single SSR measurement. **c** Frequency discrimination error as a function of total measurement time for a single SSR measurement (orange, $N_{RR} = 10^4$) and ensemble averaging (lilac, $N_{ens} = 10^4$). Each data point corresponds to a different interaction time. **d** Frequency discrimination error as a function of total measurement time, with a fixed interaction time of 44 μs ($\varphi_\Delta = \pi$). Although for SSR (orange) the total time is incremented by performing additional measurements (each using $N_{RR} = 10^4$), for ensemble averaging N_{ens} is increased. Expected errors obtained from simulations for $N_{RR} = 10^2$, 10^3, 10^4 are included (solid blue, green, and yellow lines, see Supplementary Note 4). **e** Frequency discrimination error as a function of number of trials, i.e., number of times the signal is interrogated, with a fixed interaction time of 44 μs using many SSR (each using $N_{RR} = 10^4$, orange) or ensemble averaging (lilac). Error bars are 1 SD.

time is $T_{SSR} = t_{init}^{anc} + t + N_{RR} \, t_{read}^{anc}$. For our experiments, $t_{init}^{anc} \approx 100$ μs and $t_{read}^{anc} \approx 17$ μs, so for $N_{RR} > 10$, most of the temporal overheads for SSR arise from readout of the ancilla spin via repeated mapping onto the NV electron spin (see "Methods"). For long interaction times ($\gtrsim 17 \, \mu s$), this overhead is less costly than repeating the interaction and SSR is faster than ensemble averaging, as shown in Fig. 5c.

In addition to performing only a single SSR measurement, we investigate the discrimination error when performing multiple SSR measurements. In Fig. 5d, the discrimination error is plotted as a function of the total time required by each readout strategy, including all measurement overheads. For SSR, the time is incremented by performing additional measurements, each using $N_{RR} = 10^4$ repetitive readouts, which we in the following refer to as hybrid strategy, as many individual SSR are averaged, while for ensemble averaging just N_{ens} is increased. Although about twice the number of photons is recorded for the hybrid strategy within the same time, ensemble average still performs better. The reason for this is that N_{RR} is too large; we observe that by reducing N_{RR}, the hybrid strategy can yield a smaller error as a function of time. For $N_{RR} = 10^3$, we find a slightly better scaling for the hybrid strategy using simulations, taking our measurement parameters into account (see Supplementary Note 4), as projection noise is sufficiently reduced. By further reducing to $N_{RR} = 10^2$, we find that the error exponent with the hybrid strategy is almost twice as large as the error exponent of ensemble averaging and there is a reduction of the overall error by a factor of 40 after measuring for 1 s. A more detailed analysis and optimization over N_{RR} can be found in Supplementary Note 4. In addition to benchmarking against the total measurement time, we also compare in terms of the number of trials, (i.e. the number of times the signal is queried), which becomes critical if many interrogations are prohibited, e.g. due to sample

contamination or degradation. As expected, the hybrid strategy performs significantly better, as a high confidence can be obtained from a single trial (Fig. 5e). In Fig. 5d, e the error decays exponentially with the total time (or equivalently number of queries) as expected, the error exponent for both methods is calculated explicitly in Supplementary Note 4.

It should be noted that these results apply to any problem with the aim to discriminate between two NV center states.

Optimal frequency estimation using a quantum coherent probe

A related task is to estimate a single unknown frequency ω of a signal $\mathcal{H}(\omega) = B \sin(\omega t + \theta)\sigma_Z = H(\omega)\sigma_Z$ with minimum uncertainty $\Delta\omega$. The variance $(\Delta\omega)^2$ is lower bounded by the inverse of the quantum Fisher information (QFI) $I(\omega)$[28]:

$$(\Delta\omega)^2 \geq \frac{1}{I(\omega)} \quad (5)$$

Notably, the QFI has a clear geometric meaning, as $I(\omega)$ can be written using the Bures distance d_B between quantum states:

$$I(\omega) = \lim_{\omega_\Delta \to 0} \frac{1}{\omega_\Delta^2} d_B(|\psi(\omega - \omega_\Delta/2)\rangle, |\psi(\omega + \omega_\Delta/2)\rangle)$$
$$= \lim_{\omega_\Delta \to 0} \frac{\alpha(\omega - \frac{\omega_\Delta}{2}, \omega + \frac{\omega_\Delta}{2})^2}{\omega_\Delta^2}, \quad (6)$$

where $\alpha(\omega - \frac{\omega_\Delta}{2}, \omega + \frac{\omega_\Delta}{2})$ is the angle between $|\psi(\omega \pm \omega_\Delta/2)\rangle$. Hence, maximizing the QFI is equivalent to optimizing the discrimination of $\lim_{\omega_\Delta \to 0} |\psi(\omega \pm \frac{\omega_\Delta}{2})\rangle$. As shown above, this means maximizing the accumulation of $\frac{dH}{d\omega} = \lim_{\omega_\Delta \to 0} \frac{H(\omega + \frac{\omega_\Delta}{2}) - H(\omega - \frac{\omega_\Delta}{2})}{\omega_\Delta}$, which is achieved by applying π-pulses whenever $\frac{dH}{d\omega}$ changes sign. This optimal strategy is illustrated in Fig. 6a, where π-pulses
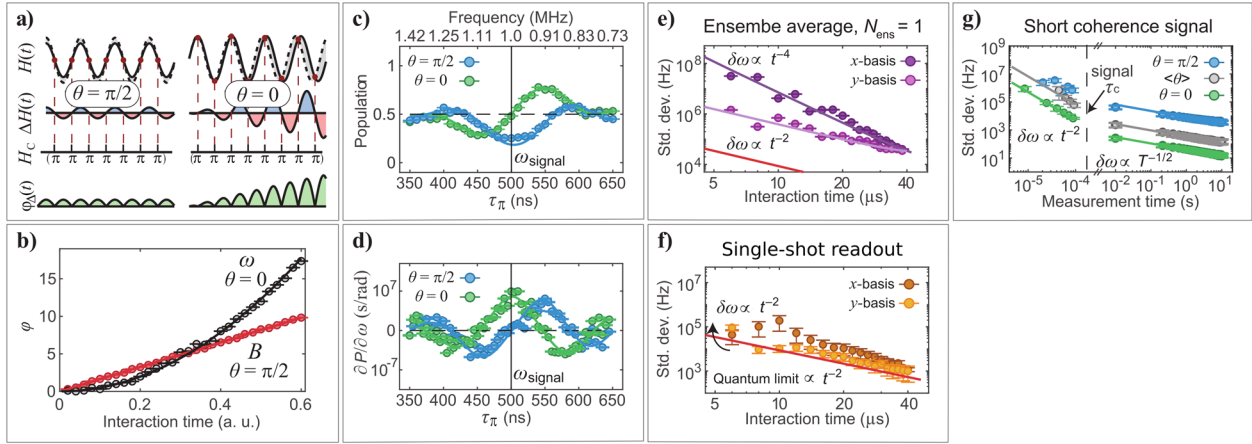
Fig. 6 Optimal quantum-frequency estimation. a Maximal sensitivity to amplitude changes is obtained for π-pulses at the signal nodes $\left(\theta = \frac{\pi}{2}\right)$, resulting in a linear accumulation of φ. Maximal sensitivity to frequency changes is obtained for π-pulses at the signal antinodes ($\theta = 0$). **b** Experimental data of the phase difference for a 1 µT change in amplitude (red) and 10 kHz change in frequency (black). Frequency estimation results in a quadratic increase of φ, whereas amplitude estimation produces a linear increase. **c** Measured NV population as a function of π-pulse spacing τ_π, given $\theta = \frac{\pi}{2}$ (blue) and $\theta = 0$ (green). Fits correspond to Supplementary Eq. 39. **d** Measured derivative of the NV population as a function of π-pulse spacing τ_π, given $\theta = \frac{\pi}{2}$ (blue) and $\theta = 0$ (green). Fits correspond to Supplementary Eq. 41. **e** Measured frequency uncertainty as a function of interaction time, for X (dark purple) or Y (light purple) measurement basis. Red line corresponds to Eq. (8) and pink line to Eq. (9). **f** Measured frequency uncertainty as a function of interaction time, for X (dark orange) or Y (light orange) measurement basis using a single SSR measurement. Red line corresponds to Eq. (8). **g** Frequency estimation for short coherence signals (for ensemble averaging): measured frequency uncertainty given $\theta = \frac{\pi}{2}$ (blue), $\theta = 0$ (green), and unknown initial phase (gray) for a single measurement with interaction time shorter than the signal coherence time (200 µs). For longer times, the sensitivity is obtained from averaging multiple measurements, each with an interaction time of 97 µs.

are applied at the antinodes of the signal with a spacing close to $\frac{\pi}{\omega}$. According to Eq. (3), it follows:

$$a\left(\omega - \frac{\omega_\Delta}{2},\ \omega + \frac{\omega_\Delta}{2}\right) = B\frac{2}{\pi}\omega_\Delta t^2 \Rightarrow I(\omega) = \left(\frac{2}{\pi}\right)^2 B^2 t^4 \tag{7}$$

The minimal uncertainty obtainable in a single experiment reads[16]:

$$\Delta\omega = \frac{\pi}{2Bt^2} \tag{8}$$

and scales again as $1/t^2$ due to the phase acceleration of the sensor. This is a special case of the analysis performed in refs. [16–19] and partially realized in ref. [15] without taking readout noise into account.

A further peculiarity is that the π-pulses should be applied close to the signal antinodes where $\frac{dH}{d\omega}$ changes sign, which is in contrast to amplitude estimation. We demonstrate the underlying intuition by recording spectra for the two edge cases—control started close to a signal node $\left(\theta = \frac{\pi}{2}\right)$ or antinode ($\theta = 0$). We again used XY8 sequences and chose the readout basis such that the final NV population reads $P = 0.5 * (1 + \sin(\varphi))$. When the π-pulses are placed at the signal nodes (Fig. 6b blue), the NV population P is maximal, but $\frac{\partial P}{\partial\omega}$ vanishes; thus, although the sensor acquires maximal phase, it is insensitive to changes in the signal frequency. In contrast, for $\theta = 0$ (Fig. 6b green), the sensor acquires minimal phase; however, small frequency changes lead to large population deviations (here, $\left(\frac{\partial P}{\partial\omega}\right)_{max} \approx 10^7\ \mathrm{Hz}^{-1}$, which corresponds to $\left(\frac{2}{4\pi}\right)Bt^2$).

Assuming perfect quantum measurements, any measurement basis in the X–Y plane would be optimal, as they all saturate the QFI (see Supplementary Note 7). The imperfect readout of the NV center not only worsens the estimate, but also provokes a dependency on the measurement basis[29]. When using conventional readout of single NV centers, the fluorescence contrast is very low and the measurement noise barely varies with the measurement basis (see Supplementary Note 8). As a consequence, it is optimal to measure (approximately) in the basis that

yields the maximal slope. In our case, assuming an initialization in the X basis, a maximal slope is obtained by measuring in the Y-basis and it follows (in the limit of small contrast, see Supplementary Note 8):

$$\Delta\omega \approx \frac{\pi}{\sqrt{8}Bt^2}\frac{\sqrt{\lambda_0 + \lambda_1}}{|\lambda_0 - \lambda_1|} \tag{9}$$

Taking our experimental values of λ_0, λ_1 into account, a factor of ~40 in SD is lost.

We realize experimental frequency estimation by measuring an unknown frequency ω (=999.3 MHz) oscillating close to 1 MHz with known amplitude (1.7 µT). Applying XY8 sequences with a pulse spacing of 500 ns and locking $\theta = 0$, we use ensemble averaging ($N_{ens} \approx 1000$) to estimate the frequency several thousand times for a variety of interaction times t (controlled by tailoring the π-pulse number). By computing the SD and assuming shot noise $\sqrt{T}$-scaling, we are able to extract an estimate for the uncertainty $\Delta\omega$ for a single measurement, which is plotted in Fig. 6c. As expected, there is a significant difference between X and Y readout (for initialization in X). Measurements in the Y basis saturate the bound in Eq. (9) and, in particular, $\Delta\omega$ scales as t^{-2}. In contrast, measuring in the X basis leads to much higher uncertainties, due to the very small slope $\frac{\partial P}{\partial\omega}$. Interestingly, using SSR (here $N_{RR} = 10^4$), the results almost coincide with the theoretical limits of perfect projective measurements (Fig. 6c lower), namely a close agreement in uncertainty using X/Y measurements and an SD of 1.6 µHz/Hz^2 within a factor of 2 to the limit, 0.9 µHz/Hz^2 set by the QFI (Eq. (8), red line in Fig. 6c).

Estimation: consequences of short coherence time

Finally, we address the relevance of this method to practical frequency estimation scenarios in NMR and communication. Although the described protocol is optimal if both, the signal and the sensor, are perfectly coherent, any deviations require further analysis of the estimation strategies.

For signals with short coherence time, which are especially relevant in nanoscale NMR where influences of molecular diffusion

and nuclear couplings limit the coherence of the NMR signal[6,30], the described protocol remains optimal, as it provides the most information per individual measurement (and correlations are not possible). As shown in Fig. 6d, the frequency can be estimated with an SD improving as t^{-2} during one sensing period, where a timing of $\theta = 0$ (green) provides three orders of magnitude improvement in SD compared to $\theta = \pi/2$ (blue). The estimate can be further improved by performing multiple measurements up to a total measurement time T, albeit with a reduced $T^{-1/2}$ scaling, typical for incoherent averages. We compute sensitivities, which are here defined as $\eta = \Delta\omega\sqrt{T}$ of 58 Hz/Hz$^{0.5}$ for $\theta = 0$ and 23 kHz/Hz$^{0.5}$ for $\theta = \pi/2$. If the signal phase is not known a priori, we cannot apply pulses at the right timing. Nevertheless, using adaptive measurements, the t^{-2} uncertainty scaling can be preserved, albeit with a higher uncertainty $\Delta\omega$. In Fig. 6d (gray), the frequency SD is plotted where the spacing τ_π was updated for each interaction time and the measurement was averaged over all signal phases (see Supplementary Note 10). Using this strategy, the frequency can be estimated with an order-of-magnitude reduction in uncertainty and sensitivity (here $\eta = 630$ Hz/Hz$^{0.5}$), but without the need to know the starting phase.

In the other limit, the coherence time of the sensor is much shorter than the coherence time of the signal. It has recently been shown that applying a chain of consecutive, correlated, measurements can enhance frequency estimation in this case[19,22,31,32]. Due to the stability of the phase of the signal, the amount of information gained from late measurements is much larger than the information gained from early ones. Hence, by employing the signal coherence, the T^4 scaling of the QFI, which relies on the coherence of the sensor, degrades to T^3. Without any control, the QFI in this regime reads (compare ref. [19]):

$$I(\omega) = \sum_\tau 4B^2 t^2 \tau^2 \, \mathrm{sinc}(\omega t/2)^2 \cos\left(\omega\tau + \frac{\omega t}{2} + \varphi\right)^2$$

$$\approx \frac{2}{3}B^2 T^3 t \, \mathrm{sinc}(\omega t/2)^2 \le 0.96 \frac{B^2 T^3}{\omega} \tag{10}$$

where the decay prefactor: $\mathrm{sinc}(\omega t/2)^2 \cos\left(\omega\tau + \frac{\omega t}{2} + \varphi\right)^2$ stems from the oscillations of the signal slope. This term can be suppressed by applying π-pulses close to the antinodes and the ultimate limit of the QFI in this regime reads (compare Supplementary Note 11):

$$I(\omega) = \left(\frac{2}{\pi}\right)^2 B^2 \sum_\tau \left((t+\tau)^2 - \tau^2\right)^2 = \frac{4}{3}\left(\frac{2}{\pi}\right)^2 B^2 T^3 t. \tag{11}$$

Hence, the scaling is the same as without control ($B^2 T^3 t$); however, there is a difference of $\sim 0.56\omega t$, which can be significant for high frequencies (taking relevant experimental values, $\omega t \sim 20$, it follows that a factor of ~10 is lost in the absence of control). Here again, the timing of the pulses is crucial. Applying resonant pulses generally leads to: $I = \frac{4}{3}\left(\frac{2}{\pi}\right)^2 B^2 T^3 t \cos(\theta)^2$ (see Supplementary Note 11). As a consequence, $I(\omega)$ vanishes for $\theta = \pi/2$ and is maximal for $\theta = 0$. More generally, applying pulses with a general timing (θ) and a general detuning δ (defined as $\omega - \pi/\tau_\pi$, where τ_π is the spacing between the π-pulses), leads to:

$$I(\omega) \approx \sum_\tau 4\left(\frac{2}{\pi}\right)^2 B^2 t^2 \tau^2 \, \mathrm{sinc}(\delta t/2)^2 \cos\left(\delta\left(\tau + \frac{t}{2}\right) + \theta\right)^2$$

$$\approx \frac{2}{3}\left(\frac{2}{\pi}\right)^2 B^2 T^3 t \begin{cases} 2\cos(\theta)^2 & \delta T \ll 1 \\ \mathrm{sinc}(\delta t/2)^2 & \delta T \gg 1 \end{cases} \tag{12}$$

It is noteworthy that in the limit of $\delta t \ll 1$, $\delta T \gg 1$ (which was implemented in ref. [19]), only a factor of 1/2 is lost, compared to the optimum, due to the detuning. Hence, in this regime, the maximal achievable advantage, compared to the results in ref. [19], is a factor of 2.

DISCUSSION

We introduced a quantum mechanical detection scheme, which achieves a quadratic increase in the sensor phase, while simultaneously reducing measurement noise, to allow frequency discrimination with a single measurement. We derived the fundamental error limit in Hamiltonian discrimination and achieved it experimentally with a suitable control. This method can provide a significant speed-up in diagnostic tests based on single quantum sensors, such as nano-NMR. In addition, we have described optimal frequency estimation strategies dependent on the readout method, and using near-ideal measurements, we obtain a frequency uncertainty near the quantum limit. If the signal phase is known, the best frequency estimate is obtained when control is started close to a signal antinode, in contradiction to methods that optimize amplitude sensitivity. These findings should prove useful in NMR, where a $\pi/2$-pulse on the sample defines the signal phase[30,31]. When the signal phase is unknown, an improved scaling can be maintained using an adaptive approach. Applications for these techniques include quantum spectroscopy and spectrum analyzers[33], characterization of quantum systems, search for dark matter, and construction of improved frequency standards.

METHODS

Experimental setup

An arbitrary waveform generator (Tektronix AWG70001A) with 20 ps timing resolution and 8 bit amplitude resolution was used for microwave control of the NV center and to generate signals of well-defined frequency and phase. Before every measurement, possible phase shifts due to different cable lengths and microwave switches and combiners were compensated for. The signal amplitude at the NV center was calibrated prior to every measurement with a standard XY8 measurement. For the fluorescence detection, we used only the first 350 ns, as we found that this duration optimizes the signal-to-noise ratio for our experimental setup.

Diamond samples

All QND measurements were performed on a diamond with 0.1% ^{13}C content[25]. NV centers in this diamond have long-phase memory times (~50 µs), while maintaining a high probability to find weakly coupled ^{13}C spins. For all other experiments, a hemispherical diamond polished into a solid immersion lens provided a higher photon detection efficiency. On the flat surface, an isotopically enriched diamond layer (99.999% ^{12}C) containing NV centers was grown by a plasma-enhanced chemical vapor deposition process, as in ref. [34]. The diamond was boiled in a 1:1:1 tri-acid mixture ($H_2SO_4 : HNO_3 : HClO_4$) for 4 h at 130 °C before experiments.

Readout statistics

For ensemble averaging, a single measurement was defined as a single interaction time followed by a single laser pulse, repeated N_{ens} times. The readout noise, δ_{SN} was determined by performing a few thousand subsequent, identical measurements and recording fluctuations from the average fluorescence intensity. For SSR statistics, a single projective measurement was performed after each interaction time. The nuclear spin state was measured with N_{RR} repetitive mapping operations which, based on the overlap of the photon intensities, results in a readout fidelity higher than 99% for $N_{RR} = 10^4$. Fluctuations in the readout state from a few thousand subsequent identical measurements was used to determine the SD, respectively.

Discrimination analysis

For y-basis readout, the probability P for the NV to remain in its initial $|0\rangle$ state is $P = \sin^2\left(\frac{\pi}{4} - \frac{\varphi}{2}\right)$ (see Fig. 2d). This formula was also used to calculate the phase accumulation as shown in Fig. 2c. For ensemble measurements, detection of a fluorescence level above or below a threshold value determined the frequency assignment. For SSR, the frequency was assigned dependent on the readout state, which was determined from a normalized fluorescence measurement (see Supplementary Note 2).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

Codes are available upon request from the authors.

REFERENCES

1. Degen, C. L., Reinhard, F. & Cappellaro, P. Quantum sensing. *Rev. Mod. Phys.* **89**, 035002 (2017).
2. Pirandola, S., Bardhan, B. R., Gehring, T., Weedbrook, C. & Lloyd, S. Advances in photonic quantum sensing. *Nat. Photonics* **12**, 724–733 (2018).
3. Wootters, W. K. Statistical distance and Hilbert space. *Phys. Rev. D* **23**, 357–362 (1981).
4. Margolus, N. & Levitin, L. B. The maximum speed of dynamical evolution. *Phys. D Nonlinear Phenom.* **120**, 188–195 (1998).
5. Helstrom, C. W. Quantum detection and estimation theory. *J. Stat. Phys.* **1**, 231–252 (1969).
6. Staudacher, T. et al. Nuclear magnetic resonance spectroscopy on a $(5\text{-nanometer})^3$ sample volume. *Science* **339**, 561–563 (2013).
7. Mamin, H. J. et al. Nanoscale nuclear magnetic resonance with a nitrogen-vacancy spin sensor. *Science* **339**, 557–560 (2013).
8. Aslam, N. et al. Nanoscale nuclear magnetic resonance with chemical resolution. *Science* **357**, 67–71 (2017).
9. Müller, C. et al. Nuclear magnetic resonance spectroscopy with single spin sensitivity. *Nat. Commun.* **5**, 4703 (2014).
10. Balasubramanian, G. et al. Nanoscale imaging magnetometry with diamond spins under ambient conditions. *Nature* **455**, 648–651 (2008).
11. Maze, J. R. et al. Nanoscale magnetic sensing with an individual electronic spin in diamond. *Nature* **455**, 644–647 (2008).
12. Sushkov, A. O. et al. Magnetic resonance detection of individual proton spins using quantum reporters. *Phys. Rev. Lett.* **113**, 197601 (2014).
13. Lovchinsky, I. et al. Nuclear magnetic resonance detection and spectroscopy of single proteins using quantum logic. *Science* **351**, 836–841 (2016).
14. Lovchinsky, I. et al. Magnetic resonance spectroscopy of an atomically thin material using a single-spin qubit. *Science* **355**, 503–507 (2017).
15. Naghiloo, M., Jordan, A. N. & Murch, K. W. Achieving optimal quantum acceleration of frequency estimation using adaptive coherent control. *Phys. Rev. Lett.* **119**, 180801 (2017).
16. Pang, S. & Jordan, A. N. Optimal adaptive control for quantum metrology with time-dependent Hamiltonians. *Nat. Commun.* **8**, 14695 (2017).
17. Gefen, T., Jelezko, F. & Retzker, A. Control methods for improved Fisher information with quantum sensing. *Phys. Rev. A* **96**, 032310 (2017).
18. Yang, J., Pang, S. & Jordan, A. N. Quantum parameter estimation with the Landau-Zener transition. *Phys. Rev. A* **96**, 020301 (2017).
19. Schmitt, S. et al. Submillihertz magnetic spectroscopy performed with a nanoscale quantum sensor. *Science* **356**, 832–837 (2017).
20. Aharon, N. et al. NV center based nano-NMR enhanced by deep learning. *Sci. Rep.* **9**, 17802 (2019).
21. Aharonov, Y., Massar, S. & Popescu, S. Measuring energy, estimating Hamiltonians, and the time-energy uncertainty relation. *Phys. Rev. A* **66**, 052107 (2002).
22. Boss, J. M., Cujia, K. S., Zopes, J. & Degen, C. L. Quantum sensing with arbitrary frequency resolution. *Science* **356**, 837–840 (2017).
23. Neumann, P. et al. Single-shot readout of a single nuclear spin. *Science* **329**, 542–544 (2010).
24. Steiner, M., Neumann, P., Beck, J., Jelezko, F. & Wrachtrup, J. Universal enhancement of the optical readout fidelity of single electron spins at nitrogen-vacancy centers in diamond. *Phys. Rev. B* **81**, 035205 (2010).
25. Unden, T. et al. Quantum metrology enhanced by repetitive quantum error correction. *Phys. Rev. Lett.* **116**, 230502 (2016).
26. Waldherr, G. et al. Dark states of single nitrogen-vacancy centers in diamond unraveled by single shot NMR. *Phys. Rev. Lett.* **106**, 157601 (2011).
27. Aslam, N., Waldherr, G., Neumann, P., Jelezko, F. & Wrachtrup, J. Photo-induced ionization dynamics of the nitrogen vacancy defect in diamond investigated by single-shot charge state detection. *N. J. Phys.* **15**, 013064 (2013).
28. Braunstein, S. L. & Caves, C. M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **72**, 3439–3443 (1994).
29. Itano, W. M. et al. Quantum projection noise: population fluctuations in two-level systems. *Phys. Rev. A* **47**, 3554–3570 (1993).
30. Schwartz, I. et al. Blueprint for nanoscale NMR. *Sci. Rep.* **9**, 6938 (2019).
31. Glenn, D. R. et al. High-resolution magnetic resonance spectroscopy using a solid-state spin sensor. *Nature* **555**, 351–354 (2018).
32. Bonato, C. et al. Optimized quantum sensing with a single electron spin using real-time adaptive measurements. *Nat. Nanotechnol.* **11**, 247–252 (2016).
33. Chipaux, M. et al. Magnetic imaging with an ensemble of nitrogen-vacancy centers in diamond. *Eur. Phys. J. D* **69**, 166 (2015).
34. Osterkamp, C. et al. Stabilizing shallow color centers in diamond created by nitrogen delta-doping using SF6 plasma treatment. *Appl. Phys. Lett.* **106**, 113109 (2015).

AUTHOR CONTRIBUTIONS

F.J., A.R., and T.G. initiated the project. S.S., L.P.M., D.L., N.S., and F.J. carried out the experiments with advice from T.G. and A.R. Analysis and interpretation of the experiments was mainly done by S.S., L.P.M., and T.G. Derivation of the theoretical principles was carried out by T.G., A.R., and L.P.M. C.O., J.L., and M.M. were responsible for sample preparation. The manuscript was written by L.P.M., S.S., and T.G. with input from all the authors.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains Supplementary Material available at https://doi.org/10.1038/s41534-021-00391-5.

Correspondence and requests for materials should be addressed to L.P.M.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.